Research Reproducibility in Political Science

Dr Nicole Janz
School of Politics and IR
Replication

Corruption

Human Rights

Political Science Replication Blog
@polscireplicate
More than half of psychology papers are not reproducible

Initiative to replicate findings of 100 prominent studies raises further questions about health of discipline

August 27, 2015

By Paul Jump   Twitter: @PaulJump
The Case of the Amazing Gay-Marriage Data: How a Graduate Student Reluctantly Uncovered a Huge Scientific Fraud

By Jesse Singal
How can we establish trust in science?
Data Access & Research Transparency (DA-RT)

- **Data transparency**: Providing full access to data itself
- **Analytic transparency**: Information about data analysis
- **Production transparency**: Process of data collection

www.dartstatement.org
Data transparency

Providing full access to data itself

Analytic transparency

Information about data analysis

Production transparency

Process of data collection

Quantitative research

Upload datasets used for analysis

Code for models (SPSS, STATA, R)

Good methods section in paper

Provide or describe raw data & variable codings

(see Lupia/Elman 2014; Moravcsik 2014)
University training
Replication exercises
Bringing the Gold Standard into the Classroom: Replication in University Teaching

Nicole Janz
University of Cambridge

Reproducibility is held to be the gold standard for scientific research. The credibility of published work depends on being able to replicate the results. However, there are few incentives to conduct replication studies in political science. Replications are difficult to conduct, time-consuming, and hard to publish because of a presumed lack of originality. This article sees a solution in a profound change in graduate teaching. Universities should introduce replications as class assignments in methods training or invest in new stand-alone replication workshops to establish a culture of replication and reproducibility. This article will
Why should you replicate?

Learn Statistics
- Real life data
- Author decisions
- Bugs included
- More fun than textbook

Reproducibility routine
- When are published results really reproducible?

Publish
- Add value
- Publish faster
Use terminology accepted in your field

Political Science (see King 2003)

<table>
<thead>
<tr>
<th>Duplication</th>
<th>Replication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verify research results</td>
<td>Test the robustness of the original research results</td>
</tr>
<tr>
<td><em>exact same data set</em></td>
<td><em>new data</em></td>
</tr>
<tr>
<td><em>exact same methods</em></td>
<td><em>new models</em></td>
</tr>
</tbody>
</table>
Practical steps in a replication study

1. Select paper
2. Access data & code
3. Identify each variable
4. Reproduce tables, figures
5. Compare

If you got to this point, you completed a duplication.
Practical steps in a replication study (II)

6 Add value
   • new data
   • new variables
   • new model specifications
   • theoretical contributions

7 Compare

8 Get feedback from peers

9 Journal submission

You now completed a full replication!
Comparing your results with the original study

Clarify with **precision** the extent to which you were able to replicate the author’s results.

Gary King (2006)

- **Exact same data and methods**: results cannot be **duplicated**

- **New** data, experiments, models, methods: describe **exactly** at which **step** the results changed and why

- **Different measures** of a concept can naturally yield different results

- **Different lab conditions** may lead to different results
Communicating failed replications

Be professional!
“We ... find that coding errors, selective exclusion of available data, and unconventional weighting of summary statistics lead to serious errors” (Herndon et al. 2013)

“If we cannot even reproduce the original results using the same publicly available data, there is no need for further commentary.” (Miller et al, 2001)
How original authors often respond

“less realistic”, “inconsistent with the substantive literature,” and “of limited utility” (Mansfield, Milner, and Rosendorff 2002)

“fundamentally flawed” (Peffley, Knigge, and Hurwitz 2001)

“statistical, computational, and reporting errors that invalidate its conclusions” (Gerber and Green 2005:301).
Our estimation approach builds off of the methodology and data used by Gomez et al. (2007) ..., adding measures of electoral closeness in order to focus on how the randomly assigned cost (rain) has a different impact depending on the electoral environment.
... we analyze a dyad-year data set (used by Rauchhaus 2009) to examine whether existing findings on the effect of symmetric nuclear weapons possession on conflict are robust to the improvements noted above. We find that once pre-nuclear dyadic conflict is controlled for, symmetric nuclear dyads are not more likely to experience low-level conflict.
Can inflation expectations be measured using commodity futures prices?

Rahould Saludeck,
Southwestern Illinois College
and
O'Marrai Cofman
Elon University

ARTICLE INFO

Abstract

This paper examines the use of commodity futures prices to measure inflation expectations for the Euro area. The analysis is based on the structural model developed by Blanchard and Kiyotaki (1996) and the commodity model developed by Jorion (1996). The results show that commodity futures prices are useful in measuring inflation expectations for the Euro area.

Keywords: Inflation, Commodity Futures Prices, Euro Area

INTRODUCTION

Inflation expectations play a crucial role in macroeconomic models. This paper examines the use of commodity futures prices to measure inflation expectations for the Euro area. The analysis is based on the structural model developed by Blanchard and Kiyotaki (1996) and the commodity model developed by Jorion (1996). The results show that commodity futures prices are useful in measuring inflation expectations for the Euro area.

Keywords: Inflation, Commodity Futures Prices, Euro Area

Questioning the Effect of Nuclear Weapons on Conflict

Mark S. Bell and Nicholas L. Miller

Abstract

We examine the effect of nuclear weapons on interstate conflict. Using a novel methodology that has been previously used, we find that the effects of nuclear weapons on interstate conflict are not statistically significant. This finding is consistent with the hypothesis that nuclear weapons have no significant effect on the likelihood of conflict.

Keywords: Nuclear Weapons, Interstate Conflict, Peace

Effect of Nuclear Weapons on Interstate Conflict

What effect do nuclear weapons have on interstate conflict? Do nuclear weapons deter the defense, deterring aggression and making states more secure? Or do nuclear weapons embolden the states that possess them, leading to conflicts that
2 Transparent workflows
4 Steps to Transparency

- Fixed Folder Structure
- Comment your code
- Clear methods section in paper
- Share materials
Fixed folder structure

Decide on a template structure for each project

Never alter raw data
FIGURE 4.1: Example Research Project File Tree

Root

ExampleProject

- Paper.Rnw
- Slideshow.Rnw
- README.md

Analysis

- Analysis1.R
- Analysis2.R

Data

- Makefile
- Gather1.R
- Gather2.R
- MergeData.R
- MainDataVariableDescriptions.md
- MainData.csv

Website.Rmd
- Main.bib
- Packages.bib
Comment your code

# loading data

# variable transformation

# merging tables

# models for table 1
Output
doedit "Z:\M14160\assessment\self data\data\stata operation-day 2-binary\EQ-I > Q-command-LR.do"

. do "C:\Users\lxwh\AppData\Local\Temp\STD02000000.tmp"

. use EQIQbinaryLR.dta,clear

. 
end of do-file

. do "C:\Users\lxwh\AppData\Local\Temp\STD02000000.tmp"

. logit BEQ2 gender representative age IQ

Iteration 0:  log likelihood = -157.28934  
Iteration 1:  log likelihood = -155.59433  
Iteration 2:  log likelihood = -155.59392  
Iteration 3:  log likelihood = -155.59392

Logistic regression                  Number of obs    =    227
                                          LR chi2(4)     =     3.39
                                          Prob > chi2    =    0.4947
Log likelihood = -155.59392            Pseudo R2       =    0.0108
# load human rights data set and select only non-OECD countries;  
# call the subset "t.non0"  
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
load("replicationdata.Rdata")  # this data file needs to be in the folder of your working directory  
t.non0 <- replicationdata[replicationdata$oecd==0, ]  # select developing nations only (non-OECD)  
t.non0 <- droplevels(t.non0)  # drop the unused levels (http://stackoverflow.com/questions/17217951/how-can-i-drop-unused-levels-from-a-data-frame)

# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
# Main models for paper
# Creating table "Total FDI and Personal Integrity Rights Protection (1983-2010)" in the main part of the paper ####
# US FDI ologit with robust standard errors clustered by country
CIRI.logit.LDV <- lrm(CIRI_PHYSINT ~ Lag_lognonnegUS_fdi_totalpGDP + Lag_logtrd + Lag_logGDPpc + Lag_logpopulation + Lag_polity2 + Lag_conf1 + Lag_CIRI_PHYSINT, data=data.CIRI.PHY.LDV, x=TRUE, y=TRUE)  
CIRI.logit.LDV.corr <- robcov(CIRI.logit.LDV, data=data.CIRI.PHY.LDV$country)  # robust errors clustered by country

PTS.logit.LDV <- lrm(PTS’ai_reversed ~ Lag_lognonnegUS_fdi_totalpGDP + Lag_logtrd + Lag_logGDPpc + Lag_logpopulation + Lag_polity2 + Lag_conf1 + Lag_PTS’ai_reversed, data=data.PTS.LDV, x=TRUE, y=TRUE)  
PTS.logit.LDV.corr <- robcov(PTS.logit.LDV, data=data.PTS.LDV$country)

# Creating "Correlation matrix (I) to (III)"
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
corstars <- function(x){  
x <- as.matrix(x)  
R <- rcorr(x, type="spearman")$r  
p <- rcorr(x)$p  
# n <- rcorr(x)$n  
mystars <- ifelse(p < .01, "***", ifelse(p < .05, "**", ifelse(p < .1, "*", " "))))  
R <- format(round(R, 3))  
Rnew <- matrix(paste(R, mystars, sep=""), ncol=ncol(x))  # removed n here
1 Clinical data

1.1 Data overview

First we load the data and R packages. The data file is part of the paper supplement, and we have also made a copy available online.

```r
<<message=FALSE,tidy=FALSE>>=
library(survival)
library(kernlab)
library(mrs); library(spatstat)
library(RColorBrewer)
library(gplots)

# load data file from local copy or from URL
if (file.exists("Schwarz2015-supplement.Rdata"){
load("Schwarz2015-supplement.Rdata")
cat("Data loaded from local copy")
} else {
load(url("http://www.markowellab.org/supplements/Schwarz2015-supplement.Rdata");
cat("Data loaded from URL")
}

The first object in the .Rdata file is a table D with patient information:

D

# Print the data table in LaTeX format for inclusion into main manuscript
library(xtable)
print(xtable(D),file="TcbleOverview.tex")
```

The first object in the .Rdata file is a table D with patient information:

<table>
<thead>
<tr>
<th>Nr</th>
<th>TH</th>
<th>CE</th>
<th>OS</th>
<th>PFS</th>
<th>dead prog</th>
<th>Hist</th>
<th>Age</th>
<th>Stage</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV03-01</td>
<td>1.7430231</td>
<td>1.2502574</td>
<td>511</td>
<td>271</td>
<td>1</td>
<td>H2OC</td>
<td>47</td>
<td>IV</td>
<td>&lt;1cm</td>
</tr>
<tr>
<td>DV03-02</td>
<td>2</td>
<td>NA</td>
<td>0.7105091</td>
<td>977</td>
<td>363</td>
<td>H5OC</td>
<td>62</td>
<td>IV</td>
<td>2cm</td>
</tr>
<tr>
<td>DV03-04</td>
<td>3.7485436</td>
<td>1.2832829</td>
<td>209</td>
<td>153</td>
<td>1</td>
<td>H3OC</td>
<td>49</td>
<td>IV</td>
<td>&gt;2cm</td>
</tr>
<tr>
<td>DV03-07</td>
<td>9</td>
<td>NA</td>
<td>0.625</td>
<td>616</td>
<td>149</td>
<td>H1OC</td>
<td>59</td>
<td>IV</td>
<td>&gt;1cm</td>
</tr>
<tr>
<td>DV03-08</td>
<td>6</td>
<td>NA</td>
<td>0.7288849</td>
<td>744</td>
<td>298</td>
<td>H1OC</td>
<td>69</td>
<td>IV</td>
<td>&gt;2cm</td>
</tr>
<tr>
<td>DV03-10</td>
<td>6.6888996</td>
<td>0.7288849</td>
<td>744</td>
<td>298</td>
<td>1</td>
<td>H2OC</td>
<td>69</td>
<td>IV</td>
<td>&gt;1cm</td>
</tr>
<tr>
<td>DV03-12</td>
<td>7.3002990</td>
<td>0.6836961</td>
<td>1887</td>
<td>358</td>
<td>1</td>
<td>NaN</td>
<td>61</td>
<td>IV</td>
<td>&gt;1cm</td>
</tr>
<tr>
<td>DV03-17</td>
<td>3</td>
<td>2.323112</td>
<td>2.126817</td>
<td>889</td>
<td>373</td>
<td>H1OC</td>
<td>51</td>
<td>IV</td>
<td>&gt;2cm</td>
</tr>
<tr>
<td>DV03-20</td>
<td>3</td>
<td>0.467828</td>
<td>0.6494353</td>
<td>1278</td>
<td>563</td>
<td>H1OC</td>
<td>71</td>
<td>IV</td>
<td>&gt;1cm</td>
</tr>
<tr>
<td>DV03-21</td>
<td>6</td>
<td>1.475638</td>
<td>1.235205</td>
<td>1139</td>
<td>303</td>
<td>H1OC</td>
<td>60</td>
<td>III</td>
<td>&gt;1cm</td>
</tr>
<tr>
<td>DV03-23</td>
<td>11</td>
<td>8.072770</td>
<td>1.843068</td>
<td>1586</td>
<td>382</td>
<td>H1OC</td>
<td>58</td>
<td>III</td>
<td>&gt;1cm</td>
</tr>
<tr>
<td>DV03-24</td>
<td>11</td>
<td>NA</td>
<td>0.6125927</td>
<td>4166</td>
<td>177</td>
<td>H1OC</td>
<td>57</td>
<td>III</td>
<td>&gt;1cm</td>
</tr>
<tr>
<td>DV03-25</td>
<td>11</td>
<td>NA</td>
<td>0.6215927</td>
<td>1166</td>
<td>177</td>
<td>H1OC</td>
<td>57</td>
<td>III</td>
<td>&gt;1cm</td>
</tr>
<tr>
<td>DV04-20</td>
<td>14</td>
<td>1.621948</td>
<td>0.6033119</td>
<td>1513</td>
<td>601</td>
<td>1</td>
<td>H3OC</td>
<td>63</td>
<td>III</td>
</tr>
<tr>
<td>DV04-21</td>
<td>15</td>
<td>0.741273</td>
<td>706</td>
<td>352</td>
<td>0</td>
<td>H1OC</td>
<td>54</td>
<td>IV</td>
<td>&gt;1cm</td>
</tr>
<tr>
<td>DV04-27</td>
<td>16</td>
<td>NA</td>
<td>4108</td>
<td>1408</td>
<td>0</td>
<td>H0OC</td>
<td>60</td>
<td>III</td>
<td>&gt;1cm</td>
</tr>
<tr>
<td>DV04-30</td>
<td>17</td>
<td>0.8551205</td>
<td>949</td>
<td>293</td>
<td>0</td>
<td>NaN</td>
<td>60</td>
<td>III</td>
<td>&gt;1cm</td>
</tr>
</tbody>
</table>
Methods section in paper

Describe methods clearly

Name exact models with citations for statistical choices

Footnote should contain software versions

If space is restricted: Appendix
Models

For the ordered categorical outcome variables, CIRI and PTS, I estimate an ordered logit model with robust standard errors clustered on country to correct for heteroskedasticity. For the continuous outcome variable, the Latent Human Rights Scores by Fariss, I employ ordinary least squares (OLS) with panel-corrected standard errors (PCSE).\textsuperscript{17} In both models, I include a lagged dependent variable ("Past") since countries that repressed their citizens in the past are more likely to use repressive acts in the future (Gurr 1988). A lagged dependent variable also corrects the serial correlation (Beck and Katz 1995, 2009).

I include a one-year lag between the outcome and predictors to allow the effect of FDI stock to spread in the country. This means that the accumulated FDI in a country in a given year is expected to correlate with rights protection in the following year, which establishes a time order and suggests a direction of causality from FDI($t-1$) to rights protection($t$).

The data set ranges from 1983 to 2010 and includes up to 121 non-OECD countries. The selection of these cases is limited to countries with available data on personal integrity rights and FDI measures (see a list of all countries in the online appendix). Two main sets of models are estimated: The first set of models includes total FDI to compare my results with previous work, while the second set of models replaces total FDI with investment in 10 industry sectors.

\textsuperscript{17} For the ordered logit models, I use the functions lrm() and robcov() from the R package “rms” Version 4.3–0, which produces the same results as the corresponding STATA command ologit with the cluster() modification (STATA Version 13.0). For OLS with PCSE, I use the functions plm() and vcovBK() from the R package “plm” Version 1.3.1, which produces the same results as the STATA command xtpcse with the pairwise specification.
Share your materials

Readme file

Data, code, variable codebook

Information to reconstruct data from original sources
Data sharing platforms

Github

UK Data Service

OSF

Inter-university Consortium for Political and Social Research

HARVARD Dataverse

CISER

CORNELL INSTITUTE for Social and Economic Research
Replication Data for: "Foreign Direct Investment and Repression: An Analysis Across Industry Sectors"  Version 1.0


Description

The impact of foreign direct investment (FDI) on repression in developing nations is still disputed. Some argue that FDI improves economic development and exports human rights values. Others criticize the exploitation of cheap labor and resources, which may lead to tensions and government oppression. Previous studies have employed aggregate FDI data with conflicting results. Alternatively, I propose that the effects depend on what kind of FDI enters a country. I build a sectoral framework to discuss how skills and technology levels, as well as the motivation for FDI, can mediate the impact. I then examine the link in a panel data analysis (1983–2010) in 121 countries, integrating sectoral FDI in several resource, manufacturing, and service industries. The results show that investment in high-skilled and high-tech sectors has positive effects. The results are robust across several measures for repression, and when accounting for sector size, regional and time effects.

Subject

Social Sciences

Keyword

human rights, foreign direct investment, repression, sectors
<table>
<thead>
<tr>
<th>Code</th>
<th>Data</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>JHR_Replication.R</td>
<td>replicationdata.tab</td>
<td>Variable_Codebook.xlsx</td>
</tr>
<tr>
<td>R Syntax - 316.3 KB - Jun 5, 2017 - 7 Downloads MD5: b38d4241da685f4d3c42b7bfbde8a486</td>
<td>Tabular Data - 1.8 MB - Jun 5, 2017 - 20 Downloads 49 Variables, 5568 Observations - UNF:6:HfNaxyecbqKGLOf0hD3cDw==</td>
<td>MS Excel (XLSX) - 38.0 KB - Jun 5, 2017 - 17 Downloads MD5: 60199b9b4e380bd8b9c647eeaac9daf1</td>
</tr>
<tr>
<td>JHR_Replication_Notes.pdf</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Replication Notes for
Foreign Direct Investment and Repression:
An Analysis Across Industry Sectors

Nicole Janz  University of Nottingham
School of Politics and International Relations
University Park
Nottingham, NG7 2RD
Email: nicole.janz@nottingham.ac.uk

September 14, 2016
Instructions for Replication

- First, load the data. This is a panel data set with country-years. Each table or figure is identified by its header as in the article or online appendix. If you are looking for a specific table, search the (admittedly long) Rscript for that particular heading.

- Run the code chunk by chunk. Many variables are created 'on the go' and re-used at a later point. Make sure to run the code in the same order, and run all the code even if you only want to replicate e.g. the last table.

- In order to preserve the largest possible sample size, I have created data sets for each model separately (based on the original table that you have loaded in R). This way, depending on the particular model, slightly different country-years are included. Make sure to create all these data sets (again, in order of the code) to run the models.
Citation

If you work with the data for your own study (replication or original work based on these data), please cite my article as well as the data set. It would be great if you could let me know about your results. A suitable citation of the data is provided by Dataverse where you downloaded the data.

Contact

Please contact me if you have any questions about the study or replication files at: nicole.janz@nottingham.ac.uk or nicolejanz@gmail.com. Any feedback on your replication attempts is more than welcome.

Acknowledgements

Gu Li and Sergio Cuesta have verified that the provided Rscript produces the tables and figures for the study. All errors remain my own.
<table>
<thead>
<tr>
<th>Data transparency</th>
<th>Analytic transparency</th>
<th>Production transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Providing full access to data itself</td>
<td>Information about data analysis</td>
<td>Process of data collection</td>
</tr>
</tbody>
</table>

**Quantitative research**

- Upload datasets used for analysis
- Code for models (SPSS, STATA, R)
- Provide or describe raw data & variable codings
- Good methods section in paper

(see Lupia/Elman 2014; Moravcsik 2014)
Replication Exercises

Transparent workflows
Five Selfish Reasons to Share Data
Quality  Establish trust & credibility in your work

Reputation  Be known as a transparent researcher in your field

Citation  Your data will be cited

Consistency  Transparent workflow makes it easier to re-use your own data later

Practicality  Meet journal & funder standards

(Markowetz 2015)
Thank you!

nicole.janz@nottingham.ac.uk

@polscireplicate

Political Science Replication Blog
<table>
<thead>
<tr>
<th>Data transparency</th>
<th>Analytic transparency</th>
<th>Production transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Providing full access to data itself</td>
<td>Information about data analysis</td>
<td>Process of data collection</td>
</tr>
</tbody>
</table>

### Qualitative research

**Provide (partial)**
- Interview transcripts
- Field notes
- Videos…

**Describe which evidence supports which claims**
- Discursive footnotes / supplement

**Explain how data were collected:**
- Interviewee selection
- Participants
- Documents…
Literature on Replication


Literature on Transparent Workflow


- Open Science Framework. Transparency and Openness Promotion (TOP) Guidelines. 
  [https://cos.io/top/](https://cos.io/top/)

- TIER Documentation Protocol
  [https://www.haverford.edu/project-tier/protocol-v2](https://www.haverford.edu/project-tier/protocol-v2)

  Retrieved from [https://osf.io/2fqnw/](https://osf.io/2fqnw/) (slides, handouts)